International Journal of Engineering Science Invention (IJESI)
ISSN (Online): 2319 – 6734, ISSN (Print): 2319 – 6726
www.ijesi.org ||Volume 7 Issue 8 Ver I || Aug 2018 || PP 60-67

Performance Improvement of Tmdm Using Hetrogenous Data Fragmentations and High Speed Clustering

[1]Mr.J.Srinivasan, [2]S.Krishna Kumar
[1]Assistant Professor, [2]M.Phil (CS) Research Scholar
[1,2]Department of Computer Science and Applications,
[1,2]Adhiparasakthi College of Arts and Science (Autonomous),G.B.Nagar, Kalavai -632506, Vellore (District)
Corresponding Author: Mr.J.Srinivasan

Abstract : Online applications including Modern Medical Informatics Services which is less demanding, faster and more affordable. The required system executes the appropriate informatics and electronics solutions efficiently for the Tele-medicine care. Notwithstanding, the increasing complexity and the fast development of the real world healthcare challenging applications make it difficult to induce the database administrative staff. We suggest three-fold approach based on data fragmentation, database websites clustering and intelligent data distribution. This approach decreases the amount of information migrated between websites during applications' execution; accomplisses cost effective communications during applications' handling and enhances applications' response time and throughput.To perform many intelligent information redistribution, we apply k-means clustering algorithms and introduce search based techniques. We characterize a data for doctor as well as patient as a concise of data set of their medical data or information from each visit. There are diverse strategies for accessing various types of medical record or reports will be provided, also we will design two web-based services, high quality data and display for many medical services. The proposed approach is approved inside by estimating the impact of utilizing our computing services' strategies on various performance features like communications cost, response time, and throughput.

Keywords - Data fragmentation; Data Distribution; K-*Means clustering, Medical database, Web Telemedicine Database Systems (WTDS).*

Date of Submission: 21-07-2018 Date of acceptance: 6-08-2018

I. INTRODUCTION

The world software applications have provoked research- rapid growth and continuous change of the realers to propose several computing services' techniques to achieve more efficient and effective management of web telemedicine database systems (WTDS).Significant research progress has been made in the past few years to improve WTDS performance. In particular, databases as a critical component of these systems have attracted many researchers. The web plays an important role in enabling healthcare services like telemedicine to serve inaccessible areas where there are few medical resources. It offers an easy and global access to patients' data without having to interact with them in person and it provides fast channels to consult specialists in emergency situations. Different kinds of patient's information such as ECG, temperature, and heart rate need to be accessed by means of various client devices in heterogeneous communications environments. WTDS enable high quality continuous delivery of patient's information wherever and whenever needed. Several benefits can be achieved by using web telemedicine services including: medical consultation delivery, transportation cost savings, data storage savings, and mobile applications support that overcome obstacles related to the performance (e.g., bandwidth, battery life, and storage), security (e.g., privacy, and reliability), and environment (e.g., scalability, heterogeneity, and availability).

The objectives of such services are to:

(i) Develop large applications that scale as the scope and workload increases,

(ii) Achieve precise control and monitoring on medical data to generate high telemedicine database system performance,

(iii) Provide large data archive of medical data records, accurate decision support systems, and trusted event-based notifications in typical clinical centers.

However, the intractable time complexity of processing large number of medical transactions and managing huge number of communications make the design of such methods a non-trivial task.

II. LITERATURE SURVEY

2.1 J.-C. Hsieh and M.-W. Hsu, "A Cloud Computing Based 12-Lead ECG Telemedicine Service,"

Multiple medical services from different web database providers may not fit the needs for improving the telemedicine database system performance. Furthermore, the services from different web database providers may not be compatible or in some cases it may increase the processing time because of the constraints on the network. Finally, there has been lack in the tools that support the design, analysis and cost-effective deployments of web telemedicine database systems.

2.2 M.T. Ozsu and P. Valduriez," Principles of Distributed Databases".

When web telemedicine database applications have contradictory requirements that avert breakdown of the relation into mutually exclusive fragments. Those applications whose views are defined on more than one fragment may suffer performance ruin. In this case, it might be necessary to retrieve data from two or more fragments and take their join, which is costly Data fragmentation technique describes how each fragment is derived from the database global relations. Three main classes of data fragmentation have been discussed in the literature; horizontal, vertical, and hybrid. Although there are various schemes describing data partitioning, few are known for the efficiency of their algorithms and the validity of their results.

2.3 K. Voges, N. Pope, and M. Brown, "Cluster Analysis of Marketing Data Examining Online Shopping Orientation: A comparison of K-means and Rough Clustering Approaches,"

A hierarchical clustering algorithm that uses similarity upper approximation derived from a tolerance (similarity) relation and based on rough set theory that does not require any prior information about the data.The presented approach results in rough clusters in which an object is a member of more than one cluster.

2.4 A. Fronczak, J. Holyst, M. Jedyank, and J. Sienkiewicz, "Higher Order Clustering Coefficients,"

Rough clustering can help researchers to discover multiple needs and interests in a session by looking at the multiple clusters that a session belongs to. However, in order to carry out rough clustering, two additional requirements, namely, an ordered value set of each attribute and a distance measure for clustering need to be specified. Clustering coefficients are needed in many approaches in order to quantify the structural network properties.

2.5 G. Mao, M. Gao, and W. Yao, "An Algorithm for Clustering XML Data Stream Using Sliding Window,"

The Clustering technique identifies groups of network sites in large web database systems and discovers better data distributions among them. This technique is considered to be an efficient method that has a major role in reducing the amount of transferred and accessed data during processing database transactions. Accordingly, clustering techniques assist us to eliminate additional communications costs between websites, thus enhancing distributed database systems performance.

2.6 Lepakshi Goud, "Achieving Availability, Elasticity and Reliability of the Data Access in Cloud Computing,"

Data distribution describes the way of allocating the disjoint fragments among the web clusters and their respective sites of the database system. This process addresses the assignment of each data fragment to the distributed database website. Data distribution related techniques aim at improving distributed database systems performance. This can be accomplished by reducing the number of database fragments that are transferred and accessed during the execution time. Additionally, Data distribution techniques attempt to increase data availability, elevate database reliability, and reduce storage overhead.

III. SYSTEM STUDY

3.1 Existing System:

- In Recent days, researchers have focused on designing web medical database management systems that satisfy specific performance levels.
- Such performance is measured by evaluating the amount of **relevant and irrelevant data** accessed and the amount of transferred medical data during transactions' processing time.
- Different techniques have been proposed to improve telemedicine database performance, optimize medical data distribution, and control medical data proliferation.
- Through the proposed techniques, high performance for such systems can be achieved by improving one of the database web management services, such as database fragmentation, data distribution, websites clustering, distributed caching, and database scalability.

3.2 Disadvantages Of Existing System:

- Some of these data records are overlapped or even redundant, which increase the I/O transactions' handling time and so the system communications overhead.
- These works mostly have investigated fragmentation, allocation and sometimes clustering issues.

- The transactions should be executed very quickly in a flexible load balancing database environment, when the number of sites in a web database system increases to a extensive scale.
- The intractable time complexity of processing large number of medical transactions and managing huge number of communications make the design of such techniques a non-trivial task.

3.3 Proposed System:

- Our approach incorporates three enhanced computing services' techniques as followed: database fragmentation, network sites clustering and fragments allocation
- We propose an estimation model to compute communications cost in order to find cost-effective data allocation solutions. We additionally perform both external and internal evaluation of our integrated approach.
- In the proposed system we develop a fragmentation computing service technique by splitting telemedicine database relations into small disjoint fragments. This method generates the minimum number of disjoint fragments need to allocated to the web servers during the data distribution phase, thus reducing the data transferred and accessed through different websites and accordingly reducing the communications cost.
- In the proposed system we present a high speed clustering service technique which groups the web telemedicine database sites into sets of clusters as for their communications cost.
- This helps in grouping the websites that are more suitable to be in one cluster to minimize data allocation operations, which in turn helps to avoid allocating redundant data.
- We propose a new computing service technique for telemedicine data allocation and redistribution services based on transactions' processing cost functions.
- To develop a user-friendly experimental tool to perform services of telemedicine data fragmentation, websites clustering, and fragments allocation, and also assist database administrators in measuring WTDS performance.
- By integrating telemedicine database fragmentation, websites clustering, and data fragments allocation into one scenario so as to ensure ultimate web telemedicine system throughput in terms of concurrency, reliability, and data availability.

3.4 Advantages Of Proposed System:

- ✓ Our integrated approach extensively improves services requirement satisfaction in web systems, and it is concluded through further investigation and experiments.
- ✓ This technique generates the minimum number of disjoint fragments that would be allocated to the web servers in the data distribution phase.
- ✓ Introduced a high speed clustering service technique which groups the web telemedicine database sites into sets of clusters as for their communications cost.

IV. DETAILED DESIGN

4.1. Data Fragmentation

With respect to fragmentation, the unit of data distribution is a vital issue. A relation is not appropriate for distribution as application views are usually subsets of relations. Therefore, the locality of applications' accesses is defined on the derivative relations subsets. Hence it is important to divide the relation into smaller data fragments and consider it for distribution over the network sites. We considered each record in each database relation as a disjoint fragment that is subject for allocation in a distributed database sites. However, large number of database fragments is generated in this method, causing a high communication cost for transmitting and processing the fragments. In contrast to this approach, we considered the whole relation as a fragment, not all the records of the fragment have to be retrieved or updated, and a selectivity matrix that indicates the percentage of accessing a fragment by a transaction is proposed. However, this research suffers from data redundancy and fragments overlapping.

4.2. Clustering Websites

Clustering service technique identifies groups of networking sites and discovers interesting distributions among large web database systems. This technique is considered as an efficient method that has a major role in reducing transferred and accessed data during transactions processing. Moreover, grouping distributed network sites into clusters helps to eliminate the extra communication costs between the sites and then enhances the distributed database system performance by minimizing the communication costs required for processing the transactions at run time. In a web database system environment where the number of sites has expanded tremendously and amount of data has increased enormously, the sites are required to manage these data and should allow data transparency to the users of the database. Moreover, to have a reliable database system, the transactions should be executed very fast in a flexible load balancing database environment. When the number of sites in a web database system increases to a large scale, the problem of supporting high system performance with consistency and availability constraints becomes crucial. Different techniques could be developed for this purpose; one of them is websites clustering.

4.3. Data Allocation (Distribution)

Data allocation describes the way of distributing the database fragments among the clusters and their respective sites in distributed database systems. This process addresses the assignment of network node(s) to each fragment. However, finding an optimal data allocation is NP-complete problem. Distributing data fragments among database websites improves database system performance by minimizing the data transferred and accessed during execution, reducing the storage overhead, and increasing availability and reliability where multiple copies of the same data are allocated.

4.4. Commercial Outsource Databases

This section investigates current commercial outsource databases and compares them with our IFCA. The commercial outsource databases support enormous data storage architecture that distributes storage and processing across several servers. It can be used to address web database system performance and scalability requirements.

4.5. Data Allocation and Replication

Data allocation techniques aim at distributing the database fragments on the web database clusters and their respective sites. We propose a heuristic fragment allocation and replication computing service to perform the processes of fragments allocation in the WTDS. Initially, all fragments are subject for allocation to all clusters that need these fragments at their sites. If the fragment authenticates positive allocation decision value i.e., allocation benefit greater than zero, for a specific cluster, then the fragment is allocated to this cluster and tested for allocation at each of its sites, if not the fragment is not allocated to this cluster. This fragment is subsequently tested for replication in each cluster of the WTDS. Accordingly, the fragment that confirms positive allocation decision value for any WTDS cluster will be allocated at that cluster and then tested for allocation at its sites. Consequently, if the fragment shows positive allocation decision value at any site of cluster that already shows positive allocation decision value, then the fragment is allocated to that site, otherwise, the fragment is not allocated. This process is repeated for all sites in each cluster that shows positive allocation decision value.

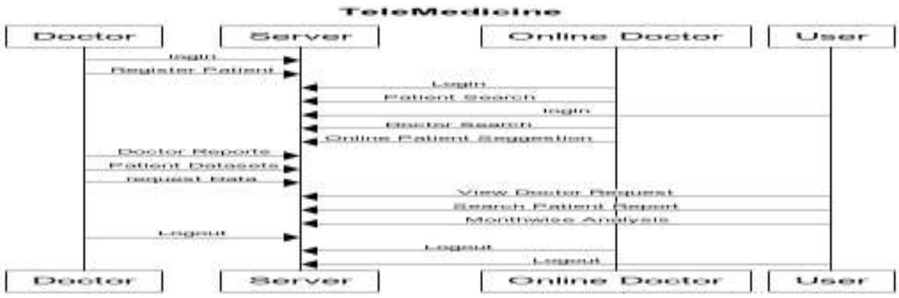

Figure 1: Telemedicine Design

- To move the medical services to the patient rather than moving patient to the medical services.
- In our project we develop a web based system in which a doctor does the patient registration and give him/her the prescription but in the case where the patient requires the special medical consultation which is available at a distance.
- The doctor send request to the online doctor about the special medical consultation and online doctor approves the request and give special consultation online.
- System also maintains patient health care records. The clustering is2 done according to the patient details, doctor details and condition of the patient.
- The fragmentation and clustering is done in order the search patient and doctor details efficiently and in real time.
- In our proposed system we introduce a fragmentation computing service technique by splitting telemedicine database relations into small disjoint fragments.
- This technique generates the minimum number of disjoint fragments that would be allocated to the web servers in the data distribution phase.
- This in turn reduces the data transferred and accessed through different websites and accordingly reduces the communications cost.
- In the proposed system we introduce a high speed clustering service technique that groups the web telemedicine database sites into sets of clusters according to their communications cost.
- This also helps in grouping the websites that are more appropriate to be in one cluster so as to minimize data allocation operations, so that allocation of redundant data is avoided.

V. ALGORITHM

5.1 K-Means Clustering

We introduce a **high speed clustering service** based on **the least average communication cost between sites.** The parameters used to control the input/output computations for generating clusters and determining the set of sites in each are computed as follows:

- Communications COST between sites CC (S_i, S_j) = data creation cost + data transmission cost between S_i, S_j.
- Communication cost range CCR (ms/byte) which is determined by telemedicine database system administrator.
- Clustering Decision Value (*cdv*)

Cluster, otherwise they are assigned to different clusters. If site Si can be assigned to more than one cluster, it will be considered for the cluster of the least average communication cost. Based on this clustering service, we develop the clustering algorithm as following:

5.2 Datafragment Algorithm

Step 1: Set 1 to I; K= F.size ()
Step 2: Do steps (3-18) until I >F.size ()
Step 3: Set I to J
Step 4: Do steps (5-16) until J> F.size ()
Step 5:If I≠J and ∃ F_i, F_j∈F go to (6)
Else, Add 1 to J and go to step (15)
Step 6: If $F_i \cap F_j \neq \emptyset$ do steps (7-14)
Else, Add 1 to J and go to step (14)
Step 7: Add 1 to K
Step 8: Create new fragment
$F_k = F_i \cap F_j$ and add it to F
Step 9: Create new fragment
$F_{k+1} = F_i - F_k$ and add it to F
Step 10: Create new fragment
$F_{k+2} = F_j - F_k$ and add it to F
Step 11: Delete F_i
Step 12: Delete F_j
Step 13: Set F+1 to J
Step 14: End IF; Step 15: End IF
Step 16: Loop
Step 17: Add 1 to I
Step 18: Loop
Step 19: Add 1 to R
Step 20: Loop

VI. Result

6.1 **Screenshot**

Figure 2 Cardiology Feature

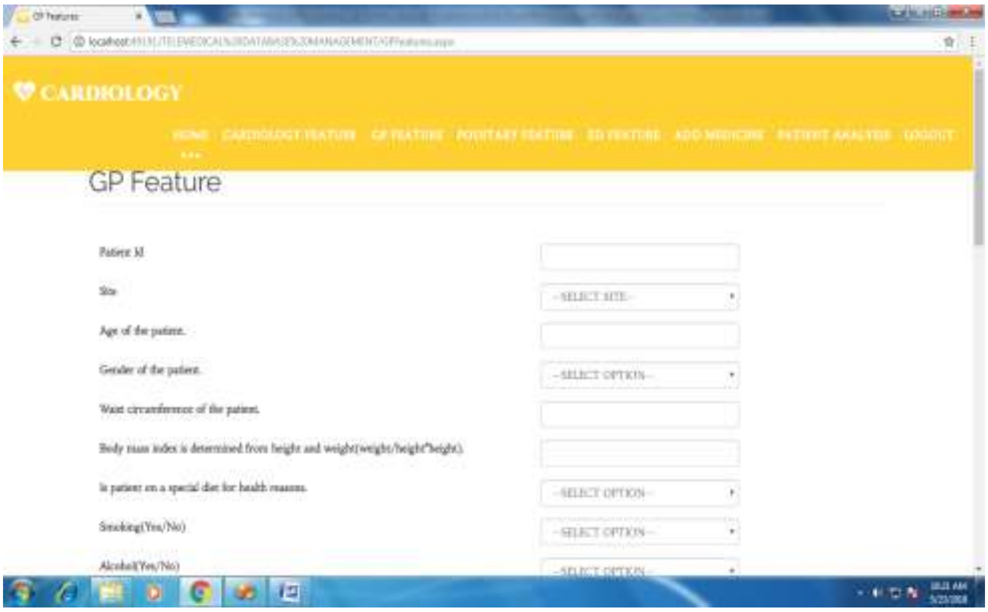

Figure 3 GP Feature

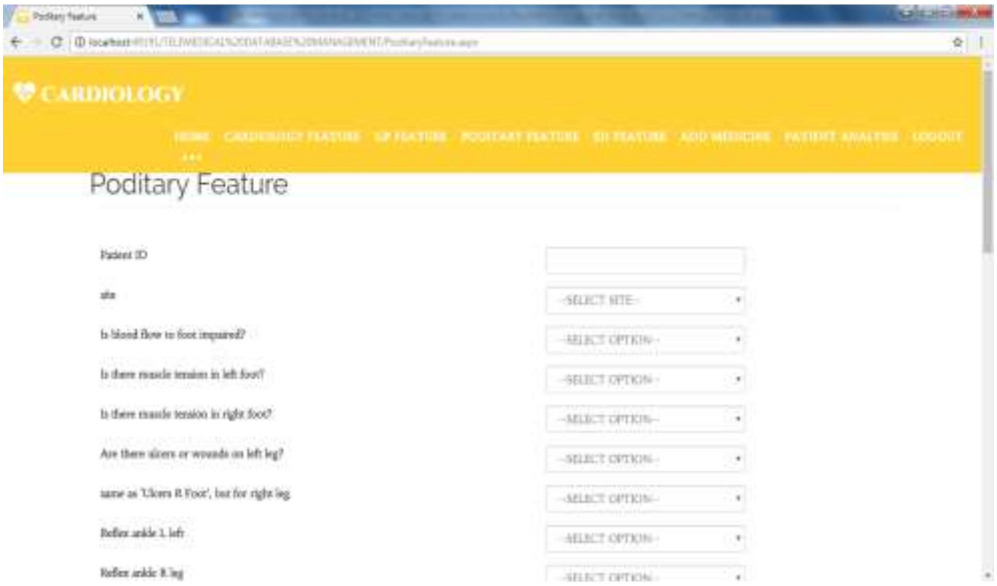

Figure 4 Podiatry Feature

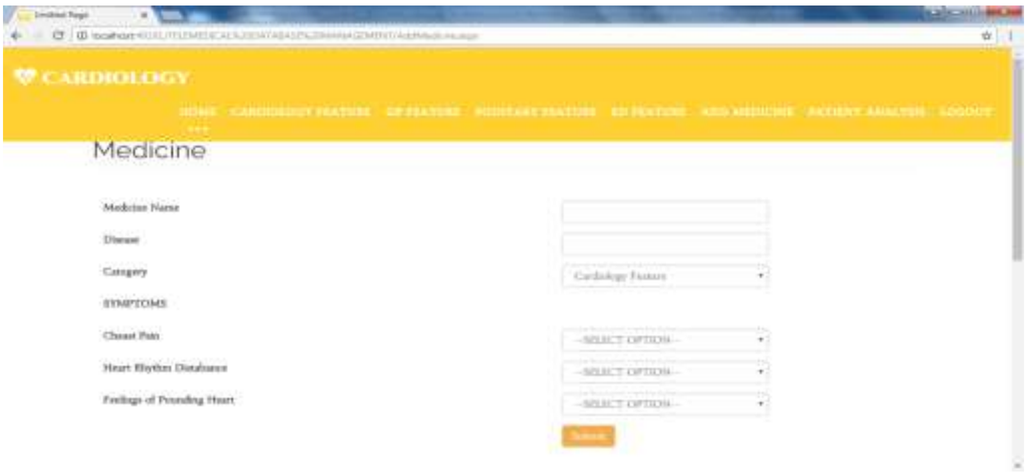

Figure 5 Add Medicine

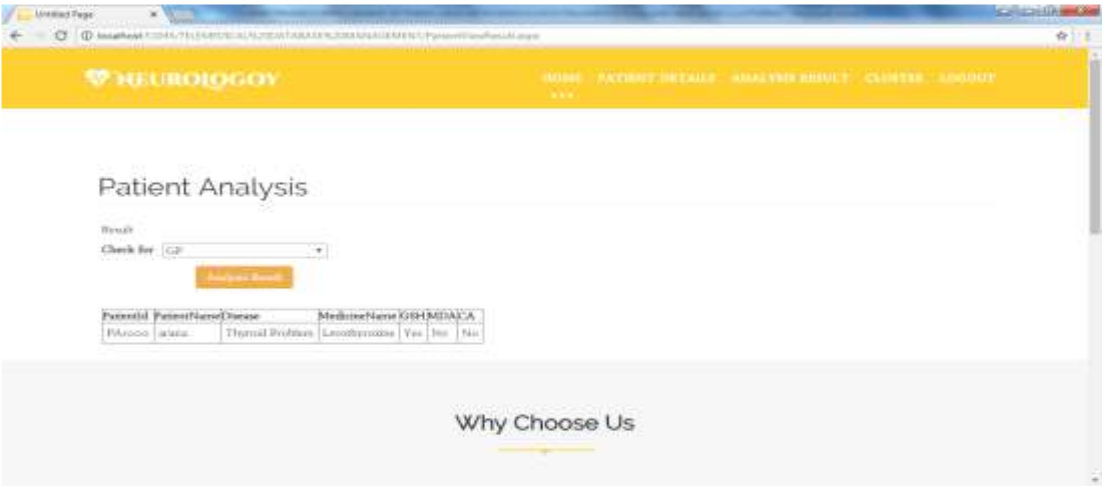

Figure 6 Patient Analysis Result

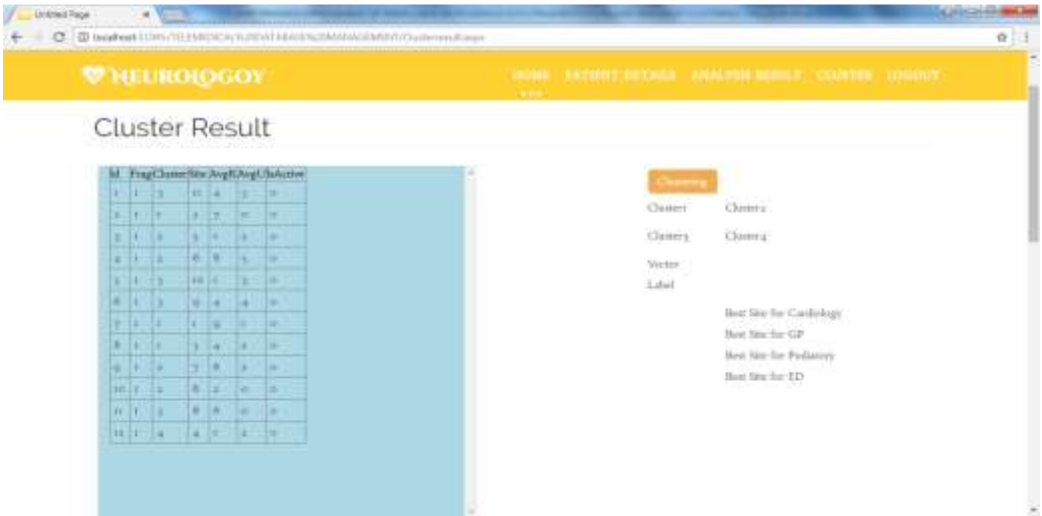

Figure 7 Clustering Results

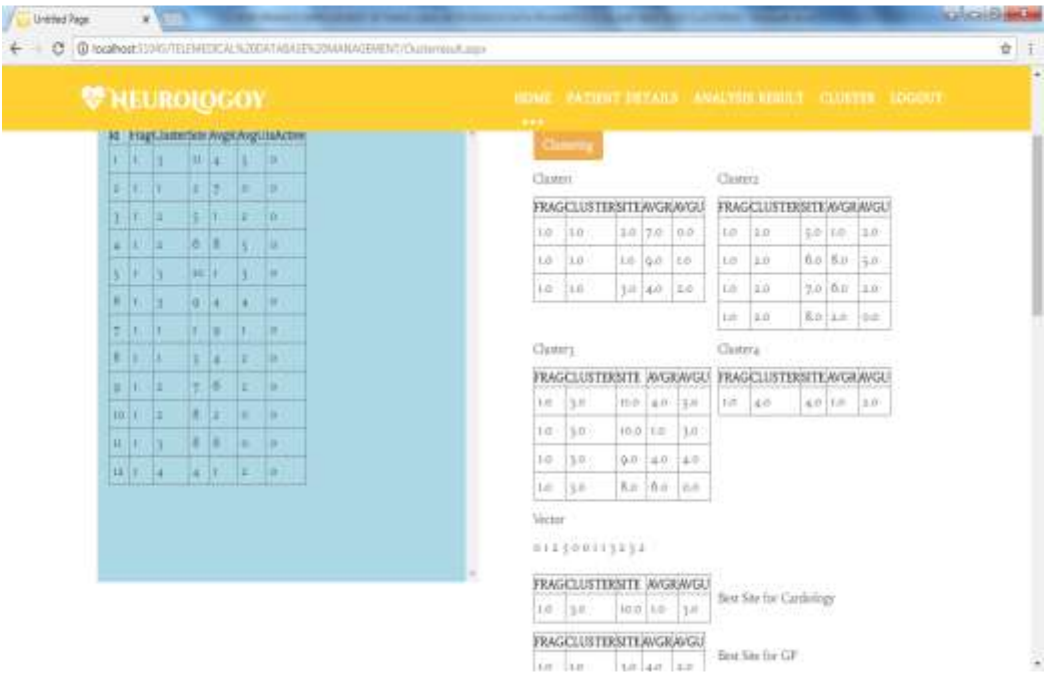

Figure 8 Results by Cluster wise

Figure 9 Best Services by Department wise

VII. Conclusion & Future Enhancement

7.1 Conclusion

In this work, we proposed a new approach to promote WTDS performance. Our approach combines three enhanced computing services' techniques namely, database fragmentation, network sites clustering and fragments allocation. We develop these techniques to solve technical challenges, like distributing data fragments among multiple web servers, handling failures, and making tradeoff between data availability and consistency. We propose an estimation model to calculate communications cost so as to assist in finding cost-effective data allocation solutions.

The novelty of our approach lies in the integration of web database sites clustering as a new component of the process of WTDS design in order to improve performance and satisfy a certain level of quality in web services. We perform both external and internal evaluation of our integrated approach. In the internal evaluation, we measure the impact of using our techniques on WTDS and web service performance measures like communications cost, response time and throughput. In the external evaluation, we compare the performance of our approach to that of other techniques in the literature. The results show that our combined approach extensively improves services requirement satisfaction in web systems. This conclusion requires more investigation and experiments.

7.2 Future Enhancement

Therefore, as future work we plan to investigate our approach on larger scale networks involving large number of sites over the cloud. We will consider applying different types of clustering and introduce search based technique to perform more intelligent data redistribution. Finally, we intend to introduce security concerns that need to be addressed over data fragments.

References

[1]. J.-C. Hsieh and M.-W. Hsu, "A Cloud Computing Based 12-Lead ECG Telemedicine Service," BMC Medical Informatics and Decision Making, vol. 12, pp. 12-77, 2012.

[2]. A. Tamhanka and S. Ram, "Database Fragmentation and Allocation: An Integrated Methodology and Case Study," IEEE Trans. Systems, Man and Cybernetics, Part A: Systems and Humans, vol. 28, no. 3, pp. 288-305, May 1998.

[3]. L. Borzemski, "Optimal Partitioning of a Distributed Relational Database for Multistage Decision-Making Support systems," Cybernetics and Systems Research, vol. 2, no. 13, pp. 809-814, 1996.

[4]. J. Son and M. Kim, "An Adaptable Vertical Partitioning Method in Distributed Systems," J. Systems and Software, vol. 73, no. 3, pp. 551-561, 2004.

[5]. S. Lim and Y. Ng, "Vertical Fragmentation and Allocation in Distributed Deductive Database Systems," J. Information Systems, vol. 22, no. 1, pp. 1-24, 1997.

[6]. S. Agrawal, V. Narasayya, and B. Yang, "Integrating Vertical and Horizontal Partitioning into Automated Physical Database Design," Proc. ACM SIGMOD Int'l Conf. Management of Data, pp. 359-370, 2004.

[7]. S. Navathe, K. Karlapalem, and R. Minyoung, "A Mixed Fragmentation Methodology for Initial Distributed Database Design," J. Computer and Software Eng., vol. 3, no. 4, pp. 395-425, 1995.

[8]. H. Ma, K. Scchewe, and Q. Wang, "Distribution Design for Higher-Order Data Models," Data and Knowledge Eng., vol. 60, pp. 400-434, 2007.

[9]. W. Yee, M. Donahoo, and S. Navathe, "A Framework for Server Data Fragment Grouping to Improve Server Scalability in Intermittently Synchronized Databases," Proc. ACM Conf. Information and Knowledge Management (CIKM), 2000.

[10]. A. Jain, M. Murty, and P. Flynn, "Data Clustering: A Review," ACM Computing Surveys, vol. 31, no. 3, pp. 264-323, 1999.

Mr.J.Srinivasan " Performance Improvement of Tmdm Using Hetrogenous Data Fragmentations and High Speed Clustering"International Journal of Engineering Science Invention (IJESI), vol. 07, no. 8, 2018, pp 60-67